

FIG. 1

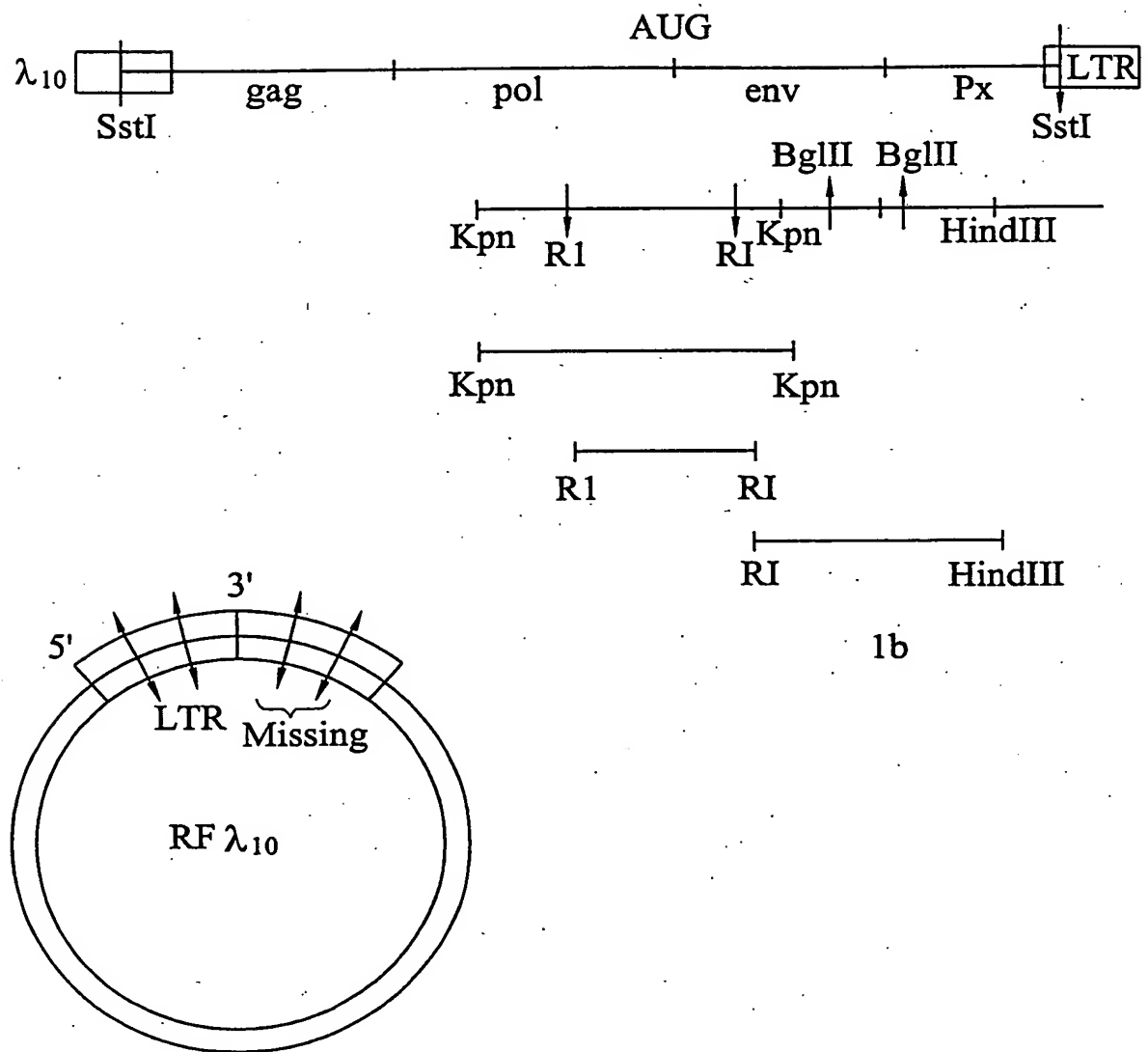
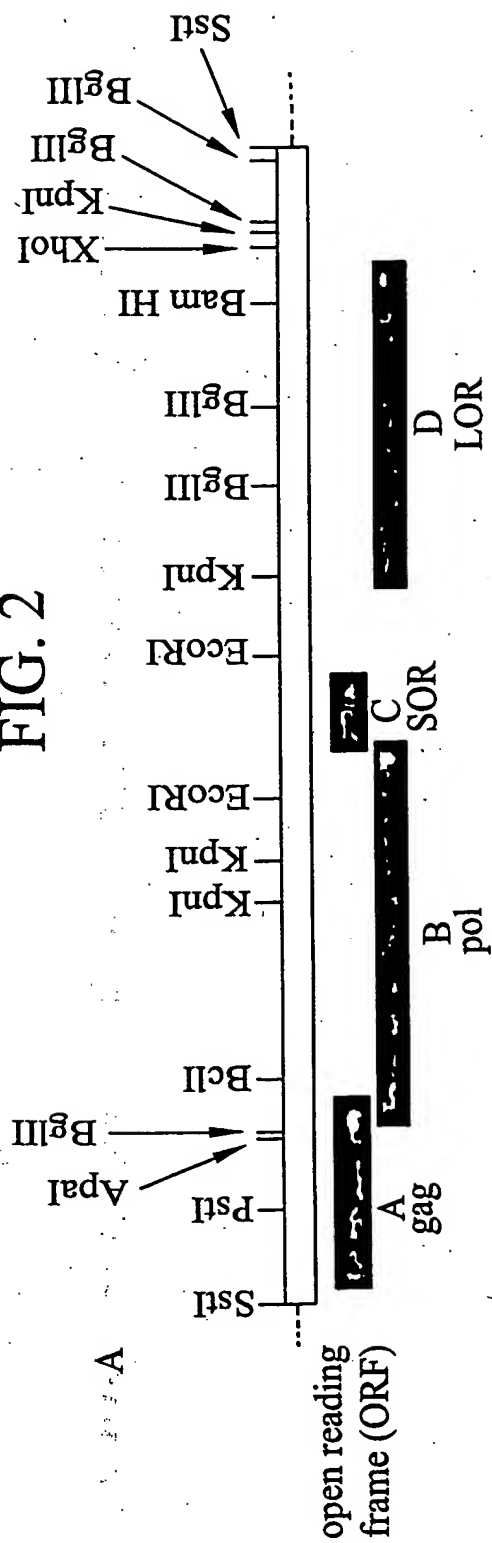
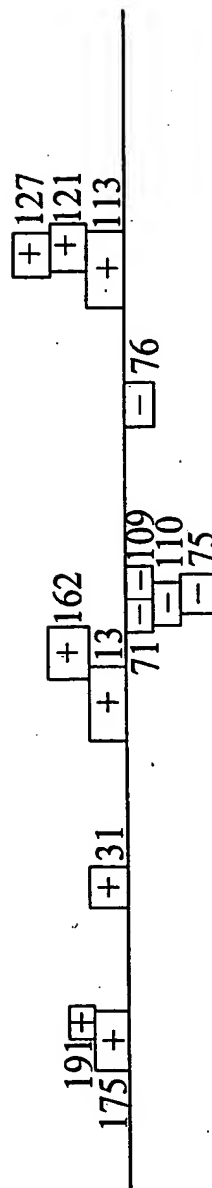


FIG. 2



**B**



BH10 BH8	(Bam HI)	IR	U3	-420
BH10 BH8				-345
BH10 BH8				-270
BH10 BH8				-195
BH10 BH8				-120
BH10 BH8				-45
BH10 BH8				-1
BH10 BH8				39
HXB2				75
HXB2				150
HXB2				221

FIG. 3 (Continued)

BH10 BH5	GAGCTCTCTCGACGAGGACTCGGCTTGCTGAAGCGGCACGGCAAGAGCGGGCGGCGACTGGTGAGTACG -----	296
BH10 BH5	CCAAAATTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGTGGGAGCGTCAGTATTAAAGCGGGGAGAATT MetGlyAlaArqAlaSerValLeuSerGlyGlyGluLeu -----	371
BH10 BH5	AGATCGATGGGAAAAATTCCGTTAAGGCCAGGGGAAAGAAAAATATAATTAAACATATAGTATGGGCAAG A pArqTrpGluLysIleArqLeuArqProGlyGlyLysLysTyrLysLeuLysHisIleValTrpAlaSer -----	446
BH10 BH5	CAGGGAGCTAGAACGATTCCGAGTTAATCCTGGCTGTAGAAACATCAGAAGGCTOTAGACAAATACTGGGACA ArqGluLeuGluArqPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCysArqGlnIleLeuGlyGln -----	521
BH10 BH5	GCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCCTCTATTGT LeuGlnProSerLeuGlnThrGlySerGluGluLeuArqSerLeuTyrAsnThrValAlaThrLeuTyrCysVal -----	596
BH10 BH5	GCATCAAAGGATAGAGATAAAAGACACCAAGGAGCTTTAGACAGATAGAGGAGAGCAAAACAAGTAAGAA HisGlnArqIleGluIleLysAspThrLysGluAlaLeuAspLysIleGluGluGluGlnAsnLysSerLysLys -----	671

Leader sequence --- GAG p17

Hind III

FIG. 3 (Continued)

BH10	AAAGCAGCAAGCAGCTGACACAGGACACAGCAGTCAGGTACGCCAAAATTACCTATAGTCAGAACAT	746	
BH5	LysAlaGlnGlnAlaAlaAspThrGlyHisSerSerGlnValSerGlnAsnTyrProIleValGlnAsnIle	138	
	-----G-----		
BH10	CCAGGGGCAATGGTACATCAGGCCCATATCACCTAGAACTTTAAATGCATGGGTAAAGTAGTAGAAGAGAGGC	821	
BH5	GlnGlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpVallysValValGluGluLysAla	163	
	-----G-----		
BH10	TTTCAGCCCAAGTAATACCCATGTTTTTCAGCATTTATCAGAAGGAGCCACCCACCAAGATTTAAACACCATGCT	896	
BH5	PheSerProGluValIleProMetPheSerAlaIleuSerGluGlyAlaThrProGlnAspLeuAsnThrMetLeu	188	
	-----G-----		
BH10	AAACACAGTGGGGGACATCAAGCAGCCATGCAAAATGTTAAAGAGAGACCATCAATGAGGAAGCTGCAGAAATGGGA	971	
BH5	AsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysGluThrIleAsnGluGluAlaAlaGluTrpAsp	213	
	-----G-----		
BH10	TAGAGTACATCCAGTGCATGCAGGGCCCTATTGCACCGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGG	1046	
BH5	ArgValHisProValHisAlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGly	238	
	-----G-----		
BH10	AACTACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCTATCCAGTAGGAGAAATTTATAA	1121	
BH5	ThrThrSerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProProIleProValGlyGluIleTyrLys	263	
	-----G-----		
BH10	AAGATGGATAATCCTGGGATTAAATAAGATGTAAGAAATGATAGCCCTACCAGCATTTCTGGACATTAAGACAAGG	1196	
BH5	ArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThrSerIleLeuAspIleArgGlnGly	288	
	-----G-----		

Pvu II

GAG p17 ---|--- GAG p26

Aha III

Aha III

Pst I

FIG. 3 (Continued)

BH10	ACCAAAGAACCTTTAGAGACTATGTAGACCGGTCTCTATAAACTCTAAGAGCCGAGCAAGCTTCACAGGAGGT	1271
BH5	ProLysGluProPheArqAspTyrValAspArqPheTyrLysThrLeuArqAlaGluGlnAlaSerGlnGluVal	313
	-----G-----C-----A-----	
	Hind III	
BH10	AAAAATTGGATGACAGAAACCTTGTGGTCCAAAATGCCAACCAGATTGTAGACACTATTTTAAAGCATTGGG	1346
BH5	LysAsnTrpMetThrGluThrLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGly	338
	-----A-----	
	Aha III	
BH10	ACCAGCGGTACACTAGAAGAAATGATGACAGCATGTTCAGGGAGTAGGAGGCCCGCCATAGGCAAGAGTTT	1421
BH5	ProAlaAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAlaArqValLeu	363
	-----	
BH10	GGCTGAAGCAATGAGCCCAAGTAACAAATACAGCTACCATTAATGATGCAGAGAGGCAATTTTAGGAACCAAGAAA	1496
BH5	AlaGluAlaMetSerGlnValThrAsnThrAlaThrIleMetMetGlnArqGlyAsnPheArqAsnGlnArqLys	388
	-----T-----A-----	
	SerThr	
BH10	GATGTTAAGTGTTCATTGTGGCAAGAAGGCGCACAGCCAGAAATTGCAGGGCCCTAGGAAAAGGGCTG	1571
BH5	MetValLysCysPheAsnCysGlyLysGluGlyHisThrAlaArqAsnCysArqAlaProArqLysLysGlyCys	413
	A---T-----I-----A-----A-----A-----A-----A-----GA-----	
	Ile Lys Arg	
	-----Direct Repeat-----POL Bql II	
BH10	TTGGAAATGTGGAAAGGACACCAAAATGAAGATTGTACTGAGAGACAGGCTAATTTTATAGGAAGATCTG	1646
BH5	TrpLysCysGlyLysGluGlyHisGlnMetLysAspCysThrGluArqGlnAlaAsnPheLeuGlyLysIleTrp	438
	-----PhePheArqGluAspLeu-----	6
	-----	
BH10	GCCTTCCTACAAGGAAGGCCAGGGAATTTTCTTCAGAGCAGACAGCCCAACAGCCCAACCATTTCTTCAGAG	1721
BH5	ProSerTyrLysGlyArqProGlyAsnPheLeuGlnSerArqProGluProThrAlaProProPheLeuGlnSer	463
	AlaPheLeuGlnGlyLysAlaArqGluPheSerSerGluGlnThrArqAlaAsnSerProThrIleSerSerGlu	31
	-----	
	-----Direct Repeat-----Direct	

FIG. 3 (Continued)

BH10	Repeat	1796	488
BH5	CAGACCCAGAGCCCAACAGCCACAGAGAGCTTCAGGCTGGGGTAGAGACAACAACCTCCCTCAGAAAGCA ArqPr GluProThrAlaProProGluGluSerPheArqSerGlyValGluThrThrThrProProGlnLysGln GlnThrArqAlaAsnSerProThrArqArqGluLeuGlnValTrpGlyArqAspAsnSerProSerGluAla ----- Ser Leu GAG p15		56
BH10	GGAGCCGATAGACAAGGAACGTATCCTTTAACTTCCCTCAGATCACTCTTGGCAACGACCCCTCGTCACAATA GluProIleAspLysGluLeuTyrProLeuThrSerLeuArqSerLeuPheGlyAsnAspProSerSerGln GlyAlaAspArqGlnGlyThrValSerPheAsnPheProGlnIleThrLeuTrpGlnArqProLeuValThrIle -----	1871	512
BH5	-----		81
BH10	AAGATAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTG LysIleGlyGlyGlnLeuLysGluAlaIleLeuAspThrGlyAlaAspAspThrValLeuGluGluMetSerLeu -----	1946	106
BH5	-----		
BH10	CCAGGAAGATGGAAACCAAAATGATAGGGGAATTGGAGGTTTTATCAAGTAAGACAGTATGATCAGATACTC Pr GlyArqTrpLysProLysMetIleGlyGlyIleGlyGlyPheIleLysValArqGlnTyrAspGlnIleLeu -----	2021	131
BH5	-----		
BH10	ATAGAAATCTGTGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT II GluIleCysGlyHisLysAlaIleGlyThrValLeuValGlyProThrProValAsnIleIleGlyArqAsn -----	2096	156
BH5	----- Aha III		
BH10	CTGTTGACTCAGATTGGTTGCACCTTTAAATTTCCCATTTAGCCCTATTGAGACGTACCAGTAAATTAAGCCA LeuLeuThrGlnIleGlyCysThrLeuAsnPheProIleSerProIleGluThrValProValLysLeuLysPro ----- A	2171	181
BH5	-----		

Position	Sequence	Position	Sequence
2246	GAAATGGATGGCCCAAAGTTAAACAATGGCCATTGACAGAGAAGAAAAATAAAAGCATTAGTAGAAATTTGTACA GlyMetAspGlyProLysValLysGlnTrpProLeuThrGluGluLysIleLysAlaLeuValGluIleCysThr	231	GAAATGGAAAAGGAAGGAAAAATTTCAAAAATTTGGCCCTGAGAAATCCATACATACTCCAGTATTTGCCATAAAG GluMetGluLysGluGlyLysIleSerLysIleGlyProGluAsnProTyrAsnThrProValPheAlaIleLys
2396	AAAAAGACAGTACTAAATGGAGAAAAATTAGTAGATTTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAA LysLysAspSerThrLysTrpArqLysLeuValAspPheArqGluLeuAsnLysArqThrGlnAspPheTrpGlu	2471	GTTCAATTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAATCAGTAACAGTACTGGATGTGGGTGATGCA ValGlnLeuGlyIleProHisProAlaGlyLeuLysLysLysSerValThrValLeuAspValGlyAspAla
2546	TATTTTTCAGTTCCTTAGATGAAGACCTTCAGGAAGTATACTGCATTTACCATACCTAGTATAAACAAATGAGACA TyrPheS rValProLeuAspGluAspPheArqLysTyrThrAlaPheThrIleProSerIleAsnAsnGluThr	2621	CCAGGGATTAGATATCAGTACAAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATG ProGlyIleArqTyrGlnTyrAsnValLeuProGlnGlyTrpLysGlySerProAlaIlePheGlnSerSerMet
2696	ACAAAAATCTTAGAGCCTTTTAAAAAACAATAATCCAGACATAGTTATCTATCAATACATGGATTTGTATGTA ThrLysIleLeuGluProPheLysLysGlnAsnProAspIleValIleTyrGlnTyrMetAspLeuTyrVal		



FIG. 3 (Continued)

BH10	GGATCTGACTTAGAAATAGGGCAGCATAGAACAAATAAGAGGAGCTGAGACAACATCTGTTGAGGTGGGACTT	2771	381
BH5	GlySerAspLeuGluIleGlyGlnHisArqThrLysIleGluGluLeuArqGlnHisLeuEuArqIrpGlyLeu		
	-----Phe		
BH10	ACCACACCAGACAAACATCAGAAAGAACCTCCATTCTTGGATGGGTTATGAACCTCCATCCTGATAAATGG	2846	406
BH5	ThrThrProAspLysLysHisGlnLysGluProProPheLeuIrpMetGlyTyrGluLeuHisProAspLysIrp		
	-----		
BH10	ACAGTACAGCCTATAGTGTGCCAGAAAGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGAAATTG	2921	431
BH5	ThrValGlnProIleValLeuProGluLysAspSerIrpThrValAsnAspIleGlnLysLeuValGlyLysLeu		
	---GA-----A		
	Pvu II		
BH10	AATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAACTCCTTAGAGGAACCAAGCAGCTA	2996	456
BH5	AsnIrpAlaSerGlnIleTyrProGlyIleLysValArqGlnLeuCysLysLeuEuArqGlyThrLysAlaLeu		
	-----I-----		
BH10	ACAGAAGTAATACCACCTAACAGAAGAGCAGAGCTAGAACTGGCAGAAACAGAGAGATTCTAAAGAACCAGTA	3071	481
BH5	ThrGluValIleProLeuThrGluGluAlaGluLeuGluLeuAlaGluAsnArqGluIleLeuLysGluProVal		
	-----		
BH10	CATGGAGTGTATTATGACCCATCAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATAT	3146	506
BH5	HisGlyValTyrTyrAspProSerLysAspLeuIleAlaGluIleGlnLysGlnGlyGlnGlyGlnIrpThrIyr		
	-----		
	Aha III		
BH10	CAAAATTTACAGAGCCATTTAAAAAATCTGAAAAACAGGAAAAATATGCAAGAAATGAGGGGTGCCACACATAATGAT	3221	531
BH5	GlnIleTyrGlnGluProPheLysAsnLeuLysThrGlyLysTyrAlaArqMetArqGlyAlaHisThrAsnAsp		
	-----		
	Aha III		
BH10	GTAAACAAATTAACAGAGGCAGTGCAAAAAATAACCACAGAAAGCATAGTAATATGGGAAAGACTCCTAAATTT	3296	556
BH5	ValLysGlnLeuThrGluAlaValGlnLysIleThrThrGluSerIleValIleIrpGlyLysThrProLysPhe		
	-----		

FIG. 3 (Continued)

BH10	AACTACCCATACAAAGGAAACATGGGAAACATGGTGACAGAGTATTGGCAAGCCACCTGGATTCTCCTGAGTGG	3371	581
BH5	LysLeuProIleGlnLysGluThrTrpGluThrTrpTrpThrGluThrTrpGlnAlaThrTrpIleProGluTrp		
	-----A-----		
BH10	GAGTTTGTTAATACCCCTCCTTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACC	3446	606
BH5	GluPheValAsnThrProProLeuValLysLeuTrpTrpGlnLeuGluLysGluProIleValGlyAlaGluThr		
	-----Kpn I-----		
BH10	TTCTATGTAGATGGGCGAGCTAACAGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAACAAAGGAAGACAA	3521	631
BH5	PheTyrValAspGlyAlaAlaAsnArqGluThrLysLeuGlyLysAlaGlyTyrValThrAsnLysGlyArqGln		
	-----G-----Ser-----Arg-----		
BH10	AAGTTGTCCCTTAACACACAAATCAGAAAAGTGAAGTACAGCAATTTATCTAGCTTTGCAGGATTCA	3596	656
BH5	LysValValProLeuThrAsnThrThrAsnGlnLysThrGluLeuGlnAlaIleTyrLeuAlaLeuGlnAspSer		
	---A---C---G---A---C---Asn-----		
	Thr His		
BH10	GGATTAGAAGTAAACATAGTAACAGACTCACAAATATGCATTAGGAATCATTCAAGCACACACAGATAAAGTGAA	3671	681
BH5	GlyL uGluValAsnIleValThrAspSerGlnTyrAlaLeuGlyIleIleGlnAlaGlnProAspLysSerGlu		
	-----I-----		
BH10	TCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAGGAAAGGTCTATCTGGCATGGGTACCAGCACAC	3746	706
BH5	SerGluLeuValAsnGlnIleIleGluGlnLeuIleLysLysGluLysValTyrLeuAlaTrpValProAlaHis		
	-----Kpn I-----		
BH10	AAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTCAGTGTGGAATCAGGAAATACTATTTTAGATGGA	3821	731
BH5	LysGlyIleGlyGlyAsnGluGlnValAspLysLeuValSerAlaGlyIleArqLysIleLeuPheLeuAspGly		
	-----		

FIG. 3 (Continued)

BH10	ATAGATAAGGCCCAAGATGAACATGAGAAATATCACAGTAAATTGGAGAGCAATGGCTAGTGTATTTTAACTGCCA	3896	756
BH5	IleAspLysAlaGlnAspGluHisGluLysTyrHisSerAsnTrpArqAlaMetAlaSerAspPheAsnLeuPro		
	-----A-----		
BH10	CCTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATOTCAGCTAAAGGAGAGCCATGCATGGACAAGTA	3971	781
BH5	PrValValAlaLysGluIleValAlaSerCysAspLysCysGlnLeuLysGlyGluAlaMetHisGlyGlnVal		
	-----Pvu II-----		
BH10	GACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAAGGAAAGTTATCCTGTTAGCAGTTTCATGTA	4046	806
BH5	AspCysSerProGlyIleTrpGlnLeuAspCysThrHisLeuGluGlyLysValIleLeuValAlaValHisVal		
	-----Aha III-----		
BH10	GCCAGTGGATATATAGAAGCAGAAAGTTATCCAGCAGAAACAGGGCAGGAAACAGCATATTTTCTTTTAAATTA	4121	831
BH5	AlaSerGlyTyrIleGluAlaGluValIleProAlaGluThrGlyGlnGluThrAlaTyrPheLeuLeuLysLeu		
	-----		
BH10	GCAGGAAGATGGCCAOTAAAAACAATACATACAGACAATGGCAGCAATTTCCACCAGTGTACGGTTAAGGCCGCC	4196	856
BH5	AlaGlyArqTrpProValLysThrIleHisThrAspAsnGlySerAsnPheThrSerAlaThrValLysAlaAla		
	-----Eco RI-----		
BH10	TOTTTGGTGGCGGGAATCAAGCAGGAATTTGGAATTCCTTACAATCCCCAAAGTCAAGCAGTAGTAGAATCTATG	4271	881
BH5	CysTrpTrpAlaGlyIleLysGlnGluPheGlyIleProTyrAsnProGlnSerGlnGlyValValGluSerMet		
	-----		
BH10	AATAAGAAATTAAGAAATATAGGACAGGTAAAGATCAGGCTGAACATCTTAAGACAGCAGTACAAAATGCCA	4346	906
BH5	AsnLysGluLeuLysLysIleIleGlyGlnValArqAspGlnAlaGluHisLeuLysThrAlaValGlnMetAla		
	-----		

FIG. 3 (Continued)

BH10	931	4421	GTATTCATCCACAATTTTAAAGAAAAGGGGGATYGGGGGTACAGTGCAGGGGAAAGAATAGTACATAATA
BH5			ValPheIleHisAsnPhelysArqlysGlyGlyIleGlyGlyTyrSerAlaGlyGluArqlleValAspIlelle
BH10	956	4496	GCAACAGACATACAAACTAAAGAAATTACAAAACAAATTACAAAATTCAAAAATTTTCGGGTTTATTACAGGGAC
BH5			AlaThra pIleGlnThrLysGluLeuGlnLysGlnIleThrLysIleGlnAsnPheArqValTyrTyrArqAsp
BH10	981	4571	AGCAGAAATCCACTTTGGAAAGGACCAGCAAGCTCCTCTGGAAAGGTGAAGGGCAGTAGTAATACAAAGATAAT
BH5			SerArqAsnProleuTrpLysGlyProAlaLysLeuLeuTrpLysGlyGlyAlaValValIleGlnAspAsn
BH10	1006	4646	AGTGACATAAAGTAGTGCCCAAGAAAGCAAGATCATATTAGGGATTATGGAAAACAGATGGCAGGTGATGAT
BH5	20		SerAspIleLysValValProArqArqlysAlaLysIleIleArqAspTyrGlyLysGlnMetAlaGlyAspAsp
			CysGlnGluGluLysGlnArqSerLeuGlyIleMetGluAsnArqTrpGlnValMetIle
BH10	1015	4721	TGTGTGGCAAGTAGACAGGATGAGGATTAGAACATGGAAAGTTAGTAAACACCATAATGATGTTTCAGGGAA
BH5	45		CysV lAlaSerArqGlnAspGluAsp
			ValTrpGlnValAspArqMetArqlleArqThrTrpLysSerLeuValLysHisMetTyrValSerGlyLys
			Arg
BH10	70	4796	AGCTAGGGGATGGTTTTATAGACATCACTATGAAAGCCCTCATCCAAGAATAAGTTCAGAAGTACACATCCCACT
BH5			AlaArqGlyTrpPheTyrArqHisTyrGluSerProHisProArqlleSerSerGluValHisIleProleu
BH10	95	4871	AGGGGATGCTAGATTGGTAATAACAACATATTGGGGTCTGCAATACAGGAGAAAGAGACTGGCATTGGGTACAGGG
BH5			GlyAspAlaArqLeuValIleThrThrTrpGlyLeuHisThrGlyGluArqAspTrpHisLeuGlyGlnGly

FIG. 3 (Continued)

BH10	AGTCTCCATAGAAATGGAGGAAAGAGATATAGCACACAAGTAGACCCCTGAACCTAGCAGACCAACTAATTCATCT	4946
BH5	ValSerIleGluTrpArqLysLysArqTyrSerThrGlnValAspProGluLeuAlaAspGlnLeuIleHisLeu	120
	-----G-----	
	Arg	
BH10	GTATTACTTTGACTGTTTTTCAGACTCTGCTATAAGAAAGGCCCTTATTAGGACACATAGTTAGCCCTAGGCTGTA	5021
BH5	TyrTyrPheAspCysPheSerAspSerAlaIleArqLysAlaLeuLeuGlyHisIleValSerProArqCysGlu	145
	-C-----T-----	
BH10	ATATCAAGCAGGACATAACAAGGTAGGATCTCTACAATACTTGGCAGCTAGCAGCATTAAATAACACCAAAAAAGAT	5096
BH5	TyrGlnAlaGlyHisAsnLysValGlySerLeuClnTyrLeuAlaLeuAlaLeuIleThrProLysLysIle	170
	-----G-----	
	Val	
BH10	AAAGCCACCTTTGCCTAGTGTACGAAACTGACAGAGGATAGATGGAACAAGCCCAAGACCAAGGCCACAG	5171
BH5	LysPr ProLeuProSerValThrLysLeuThrGluAspArqTrpAsnLysProGlnLysThrLysGlyHisArq	195
	-----A-----	
	SOR	
BH10	AGGAGCCACACAATGAATGGACACTAGAGCTTTTAGAGGAGCTTAAGAATGAAGCTGTTAGACATTTTCCTAGG	5246
BH5	GlySerHisThrMetAsnGlyHis	203
	-----A-----	
BH10	ATTTGGCTCCATGGCTTAGGGCAACATATCTATGAAACTTATGGGATACTTGGCAGGAGTGGAAGCCATAATA	5321
BH5	-----A-----	
	Eco RI	
BH10	AGAATTCTGCAACAACACTGCTGTTTATCCATTTTCAGAATTGGGTOTCGACATAGCAGAATAGCGTTACTCGACA	5396
BH5	-----A-----	

FIG. 3 (Continued)

BH10 BH5	GAGGAGAGCAAGAAATGGAGCCAGTAGATCCTAGACTAGAGCCCTGGAGCATCCAGGAAGTCAGCCTAAACTG	5471
BH10 BH5	CYTGTACCAATTGCTATTGTAAAAGTGTGCTTTCATTGCCAAGTTGTTTCATACAAAGCCTTAGGCATCT	5546
BH10 BH5 BH8	CCTATGGCAGGAAGAACGGGAGACAGCGACCGAAGACCCTCCTCAAGGCAGTCAGACTCATCAAGTTTCTCTATCAA	5621
	(Sst I) --G--A--G-A--	
BH10 BH8	AGCAGTAAGTAGTACATGTAATGCAACCTATACAAATAGCAATAGTAGTAGCAI JTTAGTAGTAGCAATAATAAGCAA	5696
BH10 BH8	TAGTTCTGTGGTCCATAGTAATCATAGAATATAGGAAATATTAGACAAAGAAATAGACAGGTTAATTGATA	5771
BH10 BH8	GACTAATAGAAAGAGCAGAGACAGTGGCAATGAGAGTGAAGGAGAAATATCAGCAGCTTGTGGAGATGGGGGTGG	5846
	ENV-LOR LysGluGlnLysThrValAlaMetArqValLysGluLysIyrGlnHisLeuTrpArqTrpGlyTrp	
BH10 BH8	AGATGGGGCACCATGCTCCTTGGGATGTTGATGATCTGTAGTGTCTACAGAAATAATTGTGGGTCACAGTCTATTAT	5921
	ArgTrpGlyThrMetLeuLeuGlyMetLeuMetIleCysSerAlaThrOulLysLeuTrpValThrValIyrIyr	
BH10 BH8	GGGTACCTGTGTGGAGGAAGCAACCACCTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTA	5996
	Kpn I GlyValProValTrpLysGluAlaThrThrLeuPheCysAlaSerAspAlaLysAlaIyrAspThrGluVal	
	Phe	

22

47

72

FIG. 3 (Continued)

BH10	CATAATGTTGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTGGTAAATGTGACA	6071
BH8	HisAsnValTrpAlaThrHisAlaCysValProThrAspProAsnProGlnGluValValLeuValAsnValThr	97
BH10	GAAATTTTAAACATGTGGAAAAATGACATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAGC	6146
BH8	GluAsnPheAsnMetTrpLysAsnAspMetValGluGlnMetHisGluAspIleIleSerLeuTrpAspGlnSer	122
BH10	CTAAGCCATGTGTAAATTAACCCCACTCTGTGTAGTTTAAAGTGCACCTGATTGGAAGATGATACTAATACC	6221
BH8	LeuLysProCysValLysLeuThrProLeuCysValSerLeuLysCysThrAspLeuLysAsnAspThrAsnThr	147
BH10	AATAGTAGCGGAGAAATGATAATGGAGAAAGGAGAGATAAAAAACCTGCTCTTCAATATCAGCACAAAGCATA	6296
BH8	AsnSerSerGlyArgMetIleMetGluLysGlyGluIleLysAsnCysSerPheAsnIleSerThrSerIle	172
	Lys	
BH10	AGAGGTAAGGTGCAGAAAGAATATGCATTTTATAACTTGATATAATACCAATAGATAATGATACTACCAGC	6371
BH8	ArgGlyLysValGlnLysGluTyrAlaPhePheTyrLysLeuAspIleIleProIleAspAsnAspThrThrSer	197
BH10	TATACGTTGACAAGTTGTAAACACCTCAGTCATTACACAGGCTGTCCAAAGGTATCCTTTGAGCCCAATTCCCATA	6446
BH8	TyrThrLeuThrSerCysAsnThrSerValIleThrGlnAlaCysProLysValSerPheGluProIleProIle	222
BH10	CATTATTGTGCCCGGCTGGTTTTCGGATTCTAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACA	6521
BH8	HisTyrCysAlaProAlaGlyPheAlaIleLeuLysCysAsnAsnLysThrPheAsnGlyThrGlyProCysThr	247
BH10	AATGTCAGCACAGTACAAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGTTAAATGGCAGTCTG	6596
BH8	AsnValSerThrValGlnCysThrHisGlyIleArgProValValSerThrGlnLeuLeuAsnGlySerLeu	272

FIG. 3 (Continued)

BH10		BqI II	GCAGAAGAGAGGTAGTAATTAGATCTGCCAATTTACAGACAATGCTAAACCATAATAGTACAGCTGAACCAA	6671
BH8		Val	AlaGluGluGlulValIleArqSerAlaAsnPheThrAspAsnAlalysThrIleIleValGlnLeuAsnGln	297
BH10		X	TCTGTAGAAAATTAATTGTACAAGACCCCAACAACAATACAAGAAAAAGTATCCGTATCCAGAGAGCAGGGAGA	322
BH8		Lys	SerValGluIleAsnCysThrArqProAsnAsnAsnThrArqlYsSerIleArqlleGlnArqGlyProGlyArq	6746
BH10		Lys	GCATTTGTTACAATAGGAAAAATAGGAAATATGAGACAAGCACATTGTAAACATTAGTAGAGCAAAAATGGAATAAC	347
BH8		Lys	AlaPheValThrIleGlyLysIleGlyAsnMetArqGlnAlaHisCysAsnIleSerArqAlalysTrpAsnAsn	6821
BH10		Aha III	ACTTTAAACAGATAGTACAAATTAAGAGAACAAATTTGGAAATAATAAAACAATAATCTTTAAGCAGTCCTCA	372
BH8		X	ThrLeuLysGlnIleAspSerLysLeuArqGluGlnPheGlyAsnAsnLysThrIleIlePheLysGlnSerSer	6896
BH10		X	GGAGGGGACCCAGAAAATTGTAACGCACAGTTTTTAATTGTGGAGGGGAAATTTTCTACTGTAAATTCACACAACCTG	397
BH8		X	GlyGlyAspProGluIleValThrHisSerPheAsnCysGlyGlyGluPhePheThrCysAsnSerThrGlnLeu	6971
BH10		X	TTTAATAGTACTTGTTTAATAGTACTTGGAGTACTAAAGGGTCAAATAACACTGAAGGAAAGTGACACAATCACC	422
BH8		{	Ph AsnSerThrTrpPheAsnSerThrTrpSerThrLysGlySerAsnAsnThrGluGlySerAspThrIleThr	7066
BH10			CTCCCATGCAGATAAAACAAATTTATAAACATGTGGCAGGAAGTAGGAAAAAGCAATGTATGCCCTCCCCATCACT	447
BH8			LeuProCysArqIleLysGlnIleIleAsnMetTrpGlnGluValGlyLysAlaMetTyrAlaProProIleSer	7121



FIG. 3 (Continued)

BH10	GGACAAATTAGATGTTTCATCAAAATATTACAGGGCTGCTATTAAACAAGAGATGGTGGTAATAGCAACAATGAGTCC	7196	472
BH8	GlyGlnIleArqCysSerSerAsnIleThrGlyLeuLeuThrArqAspGlyGlyAsnSerAsnAsnGluSer		
BH10	Bq1 II GAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAATATAAAGTAGTAAAA	7271	497
BH8	GluIlePheArqProGlyGlyGlyAspMetArqAspAsnTrpArqSerGluLeuTyrLysValValLys		
BH10	ATTGAACCATTAGGAGTAGCACCCACCAAGGCAAGAGAGAGTGGTGCAGAGAGAAAAAGACGAGTGGGAATA	7346	522
BH8	IleGluProLeuGlyValAlaProThrLysAlaLysArqValValGlnArqGluLysArqAlaValGlyIle		
BH10	GGAGCTTTGTTCTTGGGTTCTTGGAGCAGCAGGAAGCACCTATGGGCGCAGCOTCAATGACGGTGACGGTACAG	7421	547
BH8	GlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMetGlyAlaAlaSerMetThrLeuThrValGln		
BH10	GCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAATTTGCTGAGGGCTATTGAGGGCGCAACAGCATCTG	7496	572
BH8	AlaArqGlnLeuLeuSerGlyIleValGlnGlnGlnAsnLeuLeuArqAlaIleGluAlaGlnGlnHisLeu		
	-----GC-----Gly		
BH10	TTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAAATCCTGGCTGTGGAAAGATACCTAAAGGATCAA	7571	597
BH8	LeuGlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArqIleLeuAlaValGluArqTyrLeuLysAspGln		
BH10	CAGCTCCTGGGGATTGGGGTTGCTCTGGAAAACCTCATTTGCACCACCTGCTGTGCCCTTGGAAIGCTAGTTGGAGT	7646	622
BH8	GlnLeuLeuGlyIleTrpGlyCysSerGlyLysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSer		

FIG. 3 (Continued)

BH10	647	7721	7721
BH8			
BH10	672	7796	7796
BH8			
BH10	697	7871	7871
BH8			
BH10	722	7946	7946
BH8			
BH10	747	8021	8021
BH8			
BH10	772	8096	8096
BH8			
BH10	797	8171	8171
BH8			

M Hind III  
 AATAAATCTCTGGAACAGATTGGGAATAACATGACCTGGATGGAGTGGGACAGAGAAATTAAACAATTACACAAGC  
 AsnLysSerLeuGluGlnIleTrpAsnAsnMetThrTrpMetGluTrpAspArqGluIleAsnAsnTrpThrSer  
 -----

TTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAGAATGAACAAGAATTATTGGAAATTAGATAAA  
 LeuIleHisSerLeuIleGluGluSerGlnAsnGlnGlnGlnLysAsnGluGlnGluLeuGluLeuAspLys  
 -----

TGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCAATAATGATAGTAGGA  
 TrpAlaSerLeuTrpAsnTrpPheAsnIleThrAsnTrpLeuTrpTrpIleLysLeuPheIleMetIleValGly  
 -----

GGCTTGGTAGGTTTAAGAATAGTTTTTCTGTACTTCTGTAGTGAATAGAGTTAGCAGGGATATTACACCATTA  
 GlyLeuValGlyLeuArqIleValPheAlaValLeuSerValValAsnArqValArqGlnGlyTrpSerProLeu  
 -----A-----  
 Ile

TCGTTTCAGACCCACCTCCCAATCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGGTGGAGAGAGA  
 SerPheGlnThrHisLeuProIleProArqGlyProAspArqProGluGlyIleGluGluGluGlyGlyGluArq  
 -----A-----  
 Asn

GACAGAGACAGATCCATTAGTGAACGGATCCTTAGCAGCTTATCTGGGACGATCTGGGAGCCTGTGCCTC  
 AspArqAspArqSerIleArqLeuValAsnGlySerLeuAlaLeuIleTrpAspAspLeuArqSerLeuCysLeu  
 -----

TTCAGCTACCACCGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAACTTCTGGGACGCAGGGGTGG  
 PheSerTyrHisArqLeuArqAspLeuLeuIleValThrArqIleValGluLeuGlyArqArqGlyTrp  
 -----

FIG. 3 (Continued)

BH10	GAAGCCCTCAAATATTGGTGGAAATCTCCTACAGTATTGGAGTCAGGAGCTAAAGAATAGTGCTGTAGCTTGCTC	8246
BH8	GlualaleuLysTyrTrpTrpAsnLeuleuGlnTyrTrpSerGlnGluLeuLysAsnSerAlaValSerLeuleu	822
	-----A-----A-----Asn	
BH10	AATGCCACAGCTATAGCAGTAGCTAGGAGGGGACAGATAGGGTTATAGAAGTAGTACAAGGAGCTTATAGAGCTATT	8321
BH8	AsnAlaThrAlaIleAlaValAlaGluGlyThrAspArgValIleGluValValGlnGlyAlaTyrArgAlaIle	847
	-----C-----C-----Ala	
BH10	CGCCACATACCTAGAAGAATAAGACAGGGCTTGGAAAGGATTTTCTCTATAAGATGGGTGGCAAGTGGTCAAAAAG	8396
BH8	ArgHisIleProArgArgIleArgGlnGlyLeuGluArgIleLeuleu	863
	-----	
BH10	TAGTGTGGTTGGATGGCCTGCTGTAAAGGAAAGAATGAGACGAGCTGAGCCAGCAGCAGATGGGTGGGAGCAGC	8471
BH8	-----T-----	
BH10	ATCTCGAGACCTAGAAAAACATGGAGCAATCACAAATAGCAACACAGCAGCTAACAAATGCTGATTGTGCCCTGGCT	8546
BH8	-----Xh I-----T-----C-----T-----	
BH10	AGAAGCACAGAGGAGGAGGTGGTGTTCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGC	8621
BH8	-----Kpn I-----	
	PvuII BqI II Aha III U3	
BH10	AGCTGTAGATCTTAGCCACTTTTAAAGAAAGGGGACTGGAAGGGCTAATTCACCTCCCAACGAAGACAAGA	8696
BH8	-----Polypurine Tract IR-----	
BH10	TATCCTTGATCTGTGGATCTACCCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCGGGCCAGGGAT	8771
BH8	-----G-----AG-----	

FIG. 3 (Continued)

BH10 BH8	CAGATATCCACTGACCTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGAGAAGTTAGAAGAAGCCCAACAA ----- -----A-----T-----	8846
BH10 BH8	AGGAGAGAACACACAGCTTGTACACCCCTGTGAGCCCTGCATGGAATGGATGACCCGGAGAGAGAAGTGTAGAGTG ----- -----T-----	8921
BH10 BH8	GAGGTTTGACAGCCGCTAGCATTTTCATCACATGGCCCGAGAGAGCTGCATCCGGAGTACTTCAAGAAGCTGCTGACA ----- -----T-----	8996
BH10 BH8	TCGAGCTTGCTACAAGGACTTTCGGCTGGGACTTTCAGGGAGGCGTGGCCYGGCGGACTGGGGAGTGGCG ----- -----	9071
BH10 BH8	AGCCCTCAGATCCTGCATATAAGCAGCTGCTTTTGGCTGTACTGGGCTCTCTGGTTAGACCAGATCTGAGCCY ----- ----- -----Bq1 II----- -----U3-----R-----	9146
BH10 BH8	Sst I R GGGAGCTC -----	9154
HXB2	YCTGGCTAGCTAGGGAACCCACTGCTTAAGCCCTCAATAAAGCTTGGCTTGAGTGCCTCA ----- -----U5----- -----Hind III----- -----Poly(A) Sig.-----R-----	9213
HXB2	AGTAGTGTGTGCCCCGTCTGTGTGTGACTCTGGTAAGTCTGAGATCCCTCAGA ----- -----U5-----IR-----	
HXB2	CCCTTTTAGTCAGTGTGGAAATCTCTAGCA	

FIG. 4

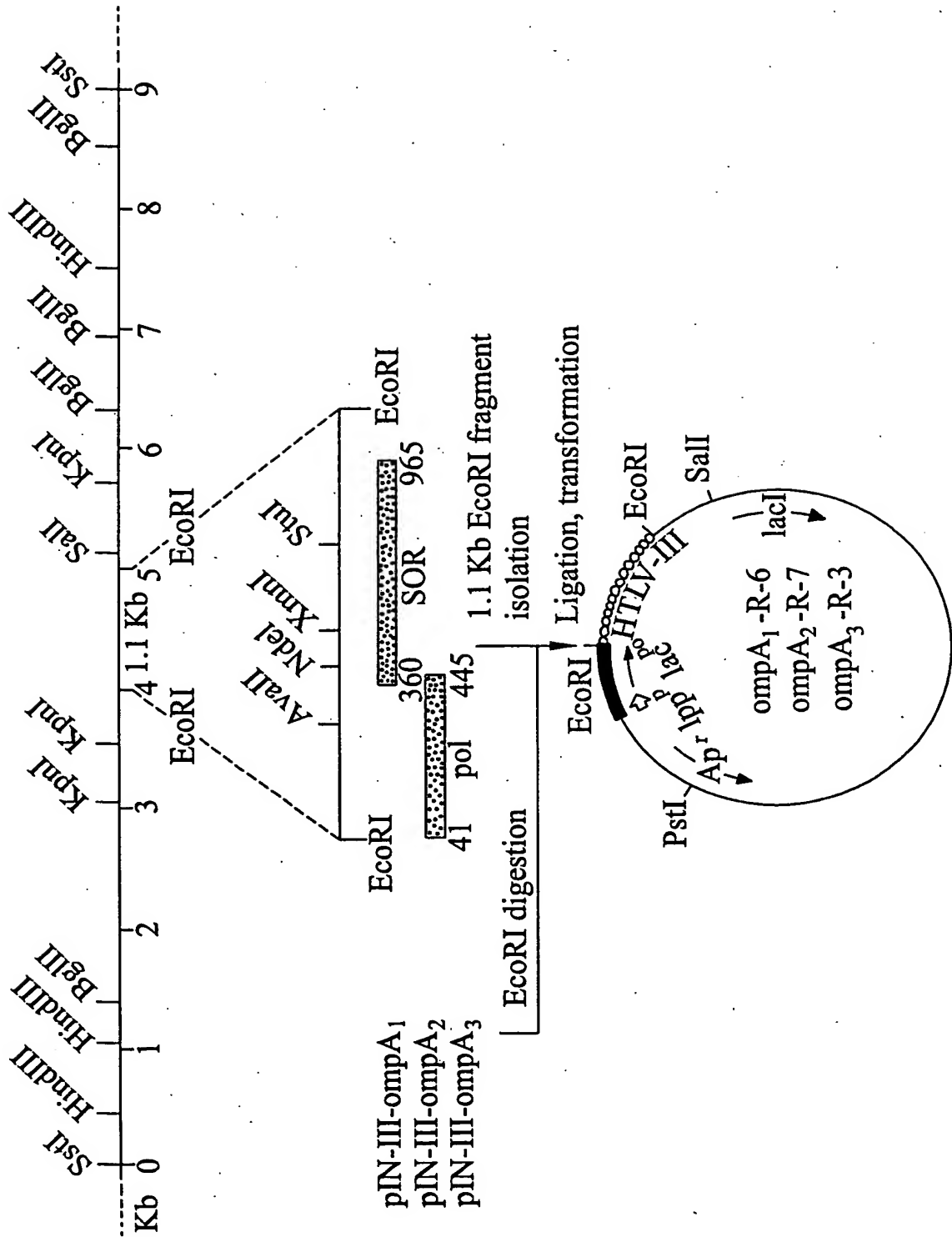


FIG. 5

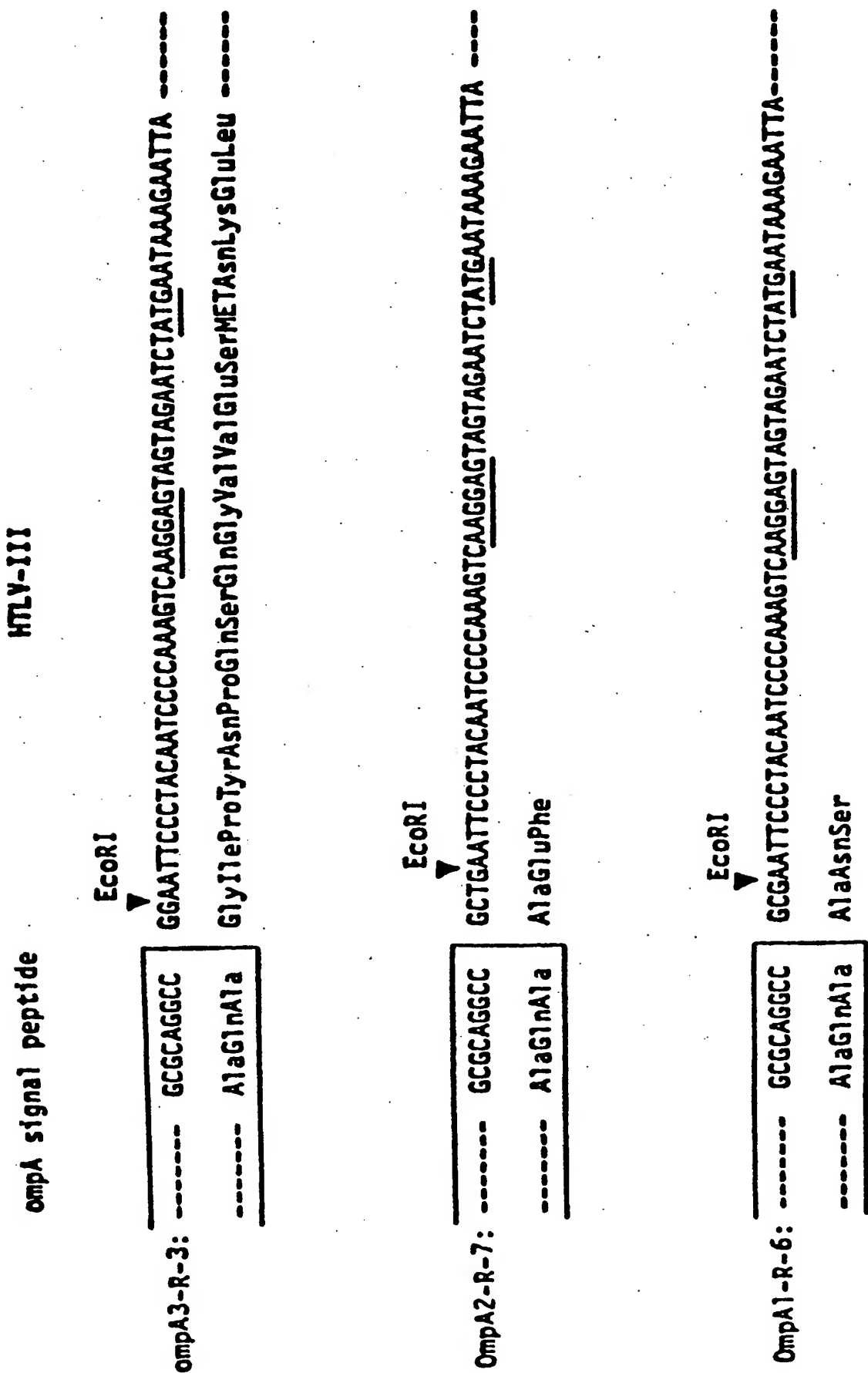


FIG. 6

